General

The Standard Two-Sided Power Distribution and its Properties: With Applications in Financial Engineering

J. René van Dorp and Samuel Kotz

This article discusses a family of distributions which would seem not to receive proper attention in the literature. The two-parameter distribution is introduced with an application in the financial engineering domain. Special cases of this family include the triangular distribution, the standard power function distribution, and the uniform distribution. Properties of the distribution are investigated and the maximum likelihood estimation procedure for its two parameters is derived. The flexibility of the family as compared to that of the beta family is discussed.

KEY WORDS: Beta distribution; Triangular distribution.

1. INTRODUCTION

We shall introduce the standard two-sided power (STSP) distribution using data on monthly interest rates for 30-year treasury maturity rates over the period from 1977–2001 (see http://www.economagic.com/popular.htm). The time series of the monthly interest rates is displayed in Figure 1(a) totaling 295 data points. Denoting the interest rate after month k by i_k, two simple financial engineering models for the random behavior of the interest rate are the additive model: $i_{k+1} = i_k + \epsilon_k$; and the multiplicative model: $i_{k+1} = i_k \nu_k$, where ϵ_k and ν_k are iid random variables (see, e.g., Leunberger 1998). Figure 1(b) depicts the time series of $\epsilon_k = i_{k+1} - i_k$. Taking logarithms in the multiplicative model yields $\log(\nu_k) = \log(i_{k+1}) - \log(i_k)$. Figure 1(c) displays the time series of $\log(\nu_k)$.

A popular model for describing the uncertainty in the observations $\log(\nu_k)$ is the normal distribution. Hence, ν_k are lognormally distributed (see, e.g., Leunberger 1998). Klein (1993) studied interest rate data on 30-year Treasury bond data from 1977–1990, finding that the empirical distribution of $\log(\nu_k)$ is "too peaky" and "fat tailed" to have been obtained from a normal distribution and rejected the lognormal hypothesis. Kozubowski and Podgórski (1999) proposed to use the asymmetric Laplace (AL) distribution with probability density function (pdf)

$$f(x|\mu,\sigma,\kappa) = \begin{cases} \frac{1}{\sigma}\frac{\kappa}{1+\kappa^2}\exp(-\frac{\sigma}{\kappa}|x-\mu|) & x < \mu \\ \frac{1}{\sigma}\frac{\kappa}{1+\kappa^2}\exp(-k\sigma|x-\mu|) & x \geq \mu, \end{cases} \quad (1)$$

where $\sigma > 0$ and $k > 0$ to capture the "peak" observed in the empirical distribution in Klein's (1993) data. Figure 2(a) displays a histogram of the differences $\log(i_{k+1}) - \log(i_k)$ for the extended dataset of 295 data points in Figure 1 (using 10 years of additional data).

Figure 2(b) provides the empirical density function of $\log(\nu_k)$ for the data in Figure 2(a), the pdf of the normal distribution and the pdf of the AL distribution, where the parameters have been estimated using maximum likelihood procedures [see, e.g., for MLE for normal distributions Johnson et al. (1994) and Kozubowski and Podgórski (1999) for MLE for AL distributions]. Both the histogram and the empirical density function exhibit the observed peak in the data. The AL distribution captures the peak in and aligns better with the empirical distribution function and may thus be considered a "better" fit. This is consistent with the conclusion in Kozubowski and Podgórski (1999).

The normal distribution has also been proposed as the uncertainty model for ϵ_k in the additive model (see, e.g., Leunberger 1998). However, since the support of a normal distribution ranges from $-\infty$ to ∞, interest rates i_k may take on negative values which causes lack of realism in the additive model (see Leunberger 1998). Figure 1(b) indicates that monthly differences of 30-year treasury bond interest rates vary in a range from -2% to 2% for the whole period 1977–2001 (with a sufficient safety margin). Hence, we propose as an uncertainty model for ϵ_k the beta distribution with support $(-0.02, 0.02)$ and pdf

$$f(x|\alpha,\beta) = \frac{1}{(0.04)^{\alpha+\beta-1}} \frac{(x+0.02)^{\alpha-1}(0.02-x)^{\beta-1}}{\mathbb{B}(\alpha,\beta)}, \quad (2)$$

$\alpha > 0$, $\beta > 0$, $\mathbb{B}(\alpha,\beta) = (\Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha+\beta))$. The beta distribution is known for its flexibility allowing a great variety of asymmetric forms (see, e.g., Johnson et al. 1995). Figure 3(a) displays the histogram of the differences $i_{k+1} - i_k$ associated with the 295 data points in Figure 1. Figure 3(b) provides the empirical density function of ϵ_k for the data in Figure 3(a). Note that, a "peak" is observed in the histogram and the empirical distribution in Figure 3(b). Figure 3(b) also displays the pdf of the beta distribution where the parameters α and β have been assessed using the MLE procedure (see, e.g., Mielke 1975). In addition, the pdf of a two-sided power distribution, to be

J. René van Dorp is Assistant Professor, and Samuel Kotz is Research Scholar and Professor, Department of Engineering Management and Systems Engineering, School of Engineering and Applied Science, The George Washington University, 1776 G Street, N.W., Washington, DC 20052 (E-mail: dorpjr@seas.gwu.edu; kotz@seas.gwu.edu). The authors are indebted to N. L. Johnson, to the editor and the associate editor of *The American Statistician*, and to the referees for their valuable comments and suggestions which improved presentation and the content of the first and second versions. In particular their insistence on a more meaningful numerical example and extending the moment ratio diagrams section enhanced our understanding of the proposed family of distributions.

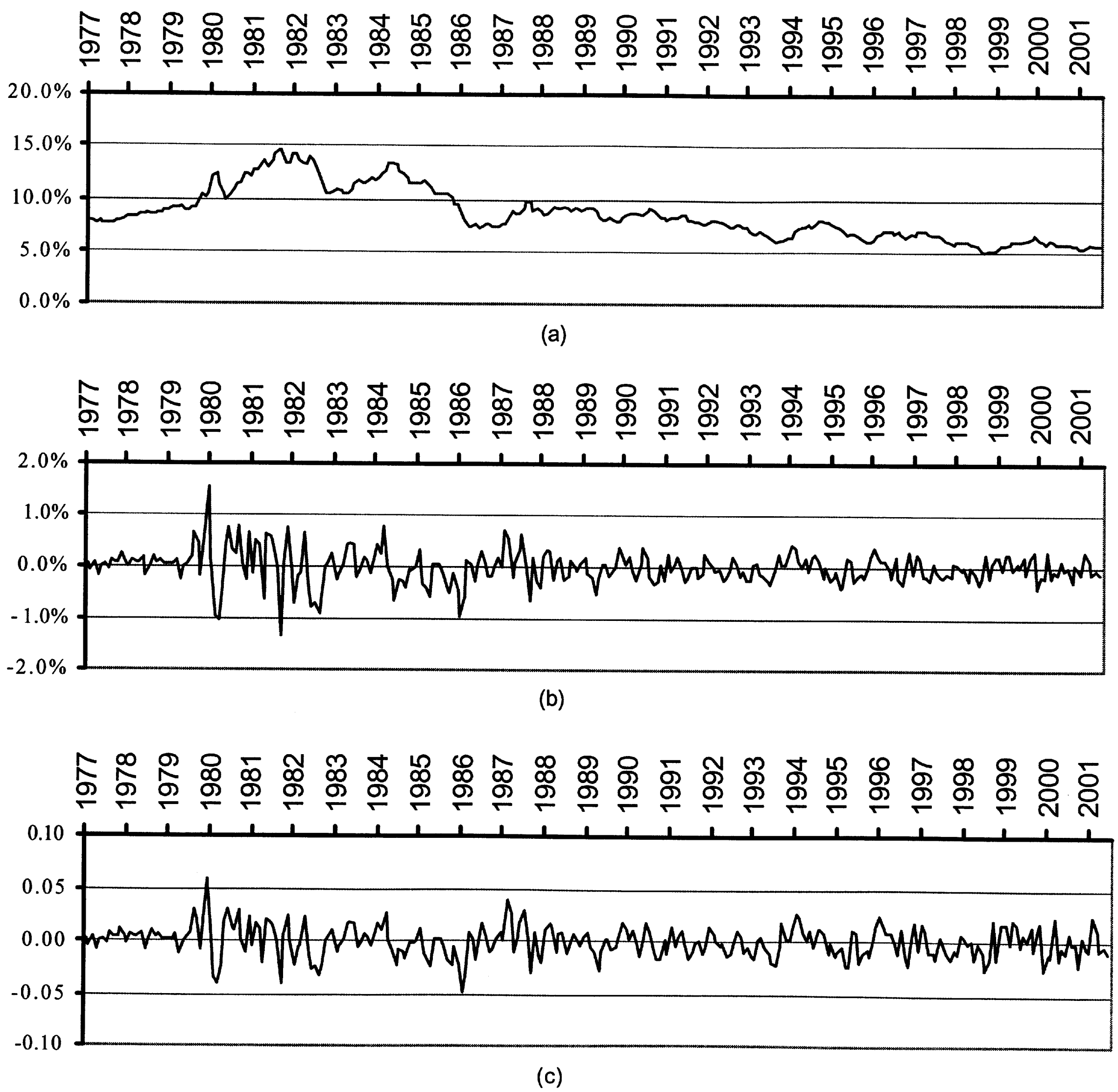

Figure 1. 30-year Monthly Treasury Bond Data from 1977–2001. (a) Interest rates; (b) monthly differences in interest rate; (c) monthly differences in log(interest rate).

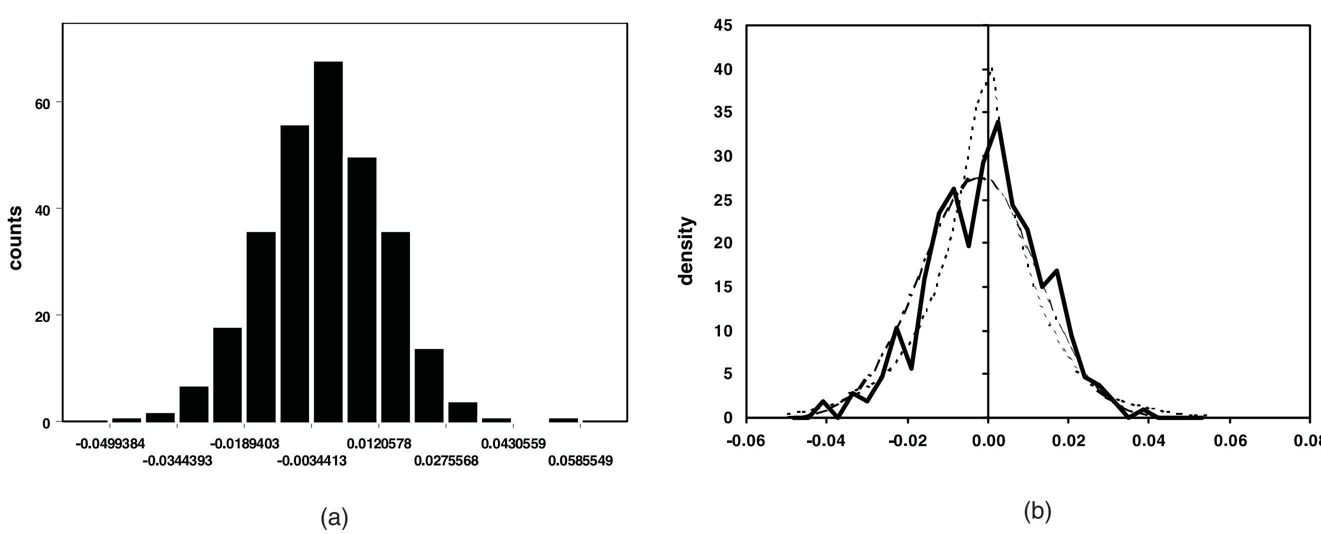

Figure 2. (a) Histogram for differences of log(interest rate)'s for the data in Figure 1(c). (b) Empirical PDF ——; Normal PDF – – – – – – ($\widehat{\mu}$ = −0.00051, $\widehat{\sigma}$ = 0.01441); asymmetric Laplace PDF - - - - - - ($\widehat{\mu}$ = −0.00051, $\widehat{\sigma}$ = 0.01121, $\widehat{\kappa}$ = 1.02309).

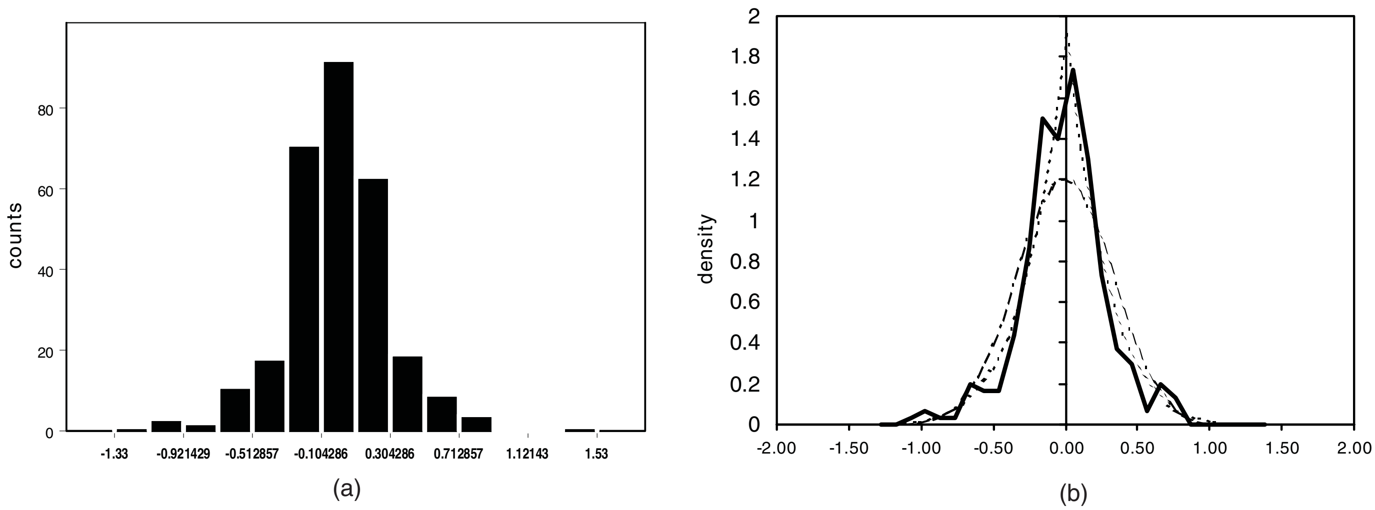

Figure 3. (a) Histogram for differences of interest rates for the data in Figure 1(b). (b) Empirical PDF ——; Beta PDF – – – – – - ($\widehat{\alpha}$ = 18.227, $\widehat{\beta}$ = 18.366); two-sided power PDF - - - - - - (see Equations (43) and (47)).

discussed in the sequel, with support $[-0.02, 0.02]$ is also presented. The two parameters of the two-sided power distribution have been estimated using the MLE procedure developed in this article. One observes from Figure 3(b) that the two-sided power distribution captures the peak in the empirical density function and aligns "better" with the empirical density function than the beta distribution. Section 4 contains additional comments concerning this comparison.

In Section 2, the standard two sided power (STSP) distribution is introduced and some basic properties are briefly discussed. In Section 3, the MLE procedure for the two parameter STSP is developed and applied to the data to the data in Figure 1(b). In Section 4, moment ratio diagrams of STSP distributions are compared with the corresponding diagrams of beta distributions. We present some brief concluding remarks in Section 5. To the best of our knowledge the STSP family is mentioned only in passing by Nadarajah (1999). We were unable to locate other literature citations.

2. STANDARD TWO-SIDED POWER DISTRIBUTIONS

Let X be a random variable with density function given by

$$f(x|\theta,n)=\begin{cases} n\left(\frac{x}{\theta}\right)^{n-1} & 0<x\leq\theta \\ n\left(\frac{1-x}{1-\theta}\right)^{n-1} & \theta\leq x<1. \end{cases} \tag{3}$$

X will be said to follow a standard two-sided power distribution $\mathrm{STSP}(\theta,n)$, $0\leq\theta\leq 1$, $n>0$, where n is not necessarily an integer. For $0\leq\theta\leq 1$ and $n>0$, the density in (3) is unimodal with the mode at θ. For $0<\theta<1$ and $0<n<1$, the form of the density function in (3) takes U-shaped forms with mode at 0 or 1. For $n=1$, the density given by (3) simplifies to the uniform$|0,1|$ density, corresponds to a triangular density on $[0,1]$ for $n=2$ and to a power function distribution for $\theta=1$. Figure 4 provides some examples of $\mathrm{STSP}(\theta,n)$ distributions, including the uniform, a triangular and a power function distribution. From (3) we obtain the cumulative distribution function (cdf) of a $\mathrm{STSP}(\theta,n)$ distribution

$$F(x|\theta,n)=\begin{cases} \theta\left(\frac{x}{\theta}\right)^{n} & 0\leq x\leq\theta \\ 1-(1-\theta)\left(\frac{1-x}{1-\theta}\right)^{n} & \theta\leq x\leq 1. \end{cases} \tag{4}$$

2.1 Moments

The kth moment of a $\mathrm{STSP}(\theta,n)$ derived from (3) is given by

$$E[X^k]=\frac{n\theta^{k+1}}{n+k}+\sum_{i=0}^{k}(-1)^i\binom{k}{k-i}\frac{n}{n+i}(1-\theta)^{i+1}. \tag{5}$$

Hence,

$$E[X] = \frac{1}{n+1}\,0+\frac{(n-1)}{n+1}\,\theta+\frac{1}{n+1}\,1=\frac{(n-1)\theta+1}{n+1}. \tag{6}$$

From (5) and (6) it follows that

$$\mathrm{var}(X)=\frac{n-2(n-1)\theta(1-\theta)}{(n+2)(n+1)^2}. \tag{7}$$

The manner in which formula (6) is written allows us to recognize the formula for the mean of the triangular distribution ($n=2$) on $[0,1]$ as a simple average of the lower bound 0, the mode θ, and the upper bound 1, from which the triangular distribution derives its intuitive appeal (see, e.g., Williams 1992). For the STSP distribution the mean is a *weighted* average of the lower bound 0, the location parameter θ, and the upper bound 1, where the weights are determined solely by n. For $n=1$, (6) simplifies to $1/2$, the mean of a uniform$[0,1]$ variable. For $n>1$, more weight is assigned to the mode θ and less to the lower (0) and upper (1) bounds as the values of n increase. In the extreme case, $n\to\infty$, no weight is assigned to the bounds and the mean simplifies to θ. For $n<1$, the mean is discounted by θ at an increasing rate as n decreases while assigning more weight to the bounds. In the extreme case , $n\downarrow 0$,the

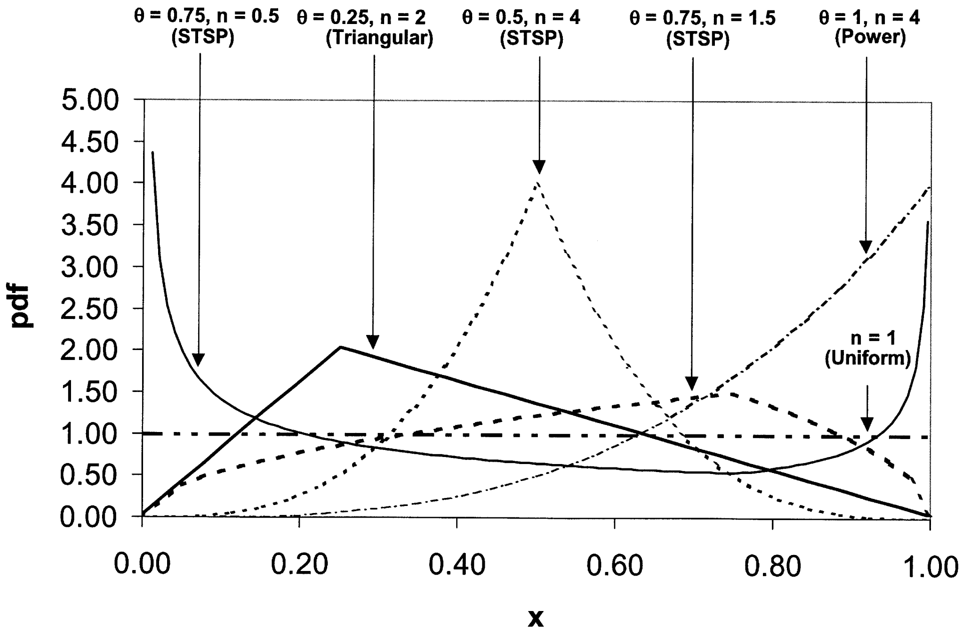

Figure 4. Some examples of STSP(θ,n) distributions.

mean simplifies to $1-\theta$. See also Section 2.4 for a discussion of the limiting distributions.

A simple relationship between the mean of a $\mathrm{STSP}(\theta, n)$ variable and its parameters (see Equation (6)) renders this family intuitive transparency, also enjoyed by the triangular distribution (see, e.g., Johnson 1997; Williams 1992).

2.2 Properties of CDF

Similarly to the beta distribution, the STSP distribution satisfies stochastically increasing and decreasing properties.

Theorem 1.
A. The cdf given by (4) is stochastically increasing for $n>1$; that is,

$$\theta_1<\theta_2,\, x\in(0,1)\Rightarrow F(x|\theta_1,n)>F(x|\theta_2,n). \quad (8)$$

B. The cdf given by (4) is stochastically decreasing for $0<n<1$; that is,

$$\theta_1<\theta_2,\, x\in(0,1)\Rightarrow F(x|\theta_1,n)<F(x|\theta_2,n). \quad (9)$$

Proof: We shall prove statement A only. (The proof of statement B is analogous.) Let $0<\theta_1<\theta_2<1,\ n>1,\ x\in(0,1)$. Three cases will be considered

$$\text{(a)}\, 0<x<\theta_1,\ \text{(b)}\, \theta_2<x<1,\quad \text{(c)}\, \theta_1<x<\theta_2. \quad (10)$$

Cases (a) and (b) are straightforward. Consider Case (c). From (4) it follows that

$$F(x|\theta_1,n)=1-\frac{(1-x)^n}{(1-\theta_1)^{n-1}},\quad F(x|\theta_2,n)=\frac{x^n}{\theta_2^{n-1}}. \quad (11)$$

For $0<\theta_1<\theta_2<1$ one derives

$$\theta_1<x<\theta_2 \Leftrightarrow \begin{cases} \frac{\theta_1}{\theta_2}<\frac{x}{\theta_2}<1 \\ 1>\frac{1-x}{1-\theta_1}>\frac{1-\theta_2}{1-\theta_1}. \end{cases} \quad (12)$$

From (12), $n>1$ and $x\in(0,1)$ we have

$$\begin{cases} 1-x<1-x\left(\frac{x}{\theta_2}\right)^{n-1} \\ (1-x)\left(\frac{1-x}{1-\theta_1}\right)^{n-1}<(1-x). \end{cases} \quad (13)$$

Hence,

$$\frac{(1-x)^n}{(1-\theta_1)^{n-1}}\ <\ 1-\frac{x^n}{\theta_2^{n-1}} \Leftrightarrow 1-\frac{(1-x)^n}{(1-\theta_1)^{n-1}}>\frac{x^n}{\theta_2^{n-1}}, \quad (14)$$

which proves statement A.

2.3 Quantile Properties

Denote by x_p the pth percentile—that is, $F(x_p|\theta,n)=p$. Properties 2, 3, and 4 below may be used for an initial estimation of the parameter θ. Properties 3 and 4 involve two somewhat unexpected relations satisfied by its quantiles x_p and x_{1-p}.

Property 1: From the cdf (4) it follows that: $x_p<p \Leftrightarrow p<\theta$.

Property 2: Analogously, $x_p=\theta \Leftrightarrow p=\theta$, regardless of the value of n. Hence, for all STSP distributions the probability mass is split at θ into θ and $(1-\theta)$.

Property 3: Consider values of p such that $p<\min(\theta,1-\theta)$. From (4) and Property 1, dealing separately with the dual definition of $F(\cdot|\theta,n)$, we have

$$\frac{x_p^n}{\theta^{n-1}}\ =\ \frac{(1-x_{1-p})^n}{(1-\theta)^{n-1}} \Leftrightarrow \frac{x_p}{1-x_{1-p}}=\left(\frac{\theta}{1-\theta}\right)^{\frac{n-1}{n}}. \quad (15)$$

Hence the ratio $x_p/1-x_{1-p}$ does not depend on p for $p<\min(\theta,1-\theta)$.

Property 4: Analogously, for $p > \max(\theta, 1-\theta)$

$$\frac{x_{1-p}}{1-x_p} = \left(\frac{\theta}{1-\theta}\right)^{\frac{n-1}{n}} \tag{16}$$

is independent of p.

2.4 Limiting Distributions

Let $X \sim \mathrm{STSP}(\theta, n)$. Setting $\theta = 0$ and letting $n \to \infty$ it follows from (6) and (7) that the distribution of X converges to a degenerate distribution with a point mass of 1 at 0. Analogously, setting $\theta = 1$ as $n \to \infty$ the distribution of X converges to a degenerate distribution of 1 at 1. Setting $0 < \theta < 1$, and letting $n \to \infty$ it follows that the distribution of X converges to a degenerate distribution with a point mass of 1 at θ. As $n \downarrow 0$ and $0 < \theta < 1$ we obtain from (5) that $\lim_{n \downarrow 0} E[X^k] = (1-\theta)$ for all k. Hence, when $n \downarrow 0$ $E[X^k]$ converges to the moments of a Bernoulli distribution with a point mass of $1-\theta$ at 1. As both X and a Bernoulli variable have bounded support, it thus follows by the uniqueness theorem for distributions with a bounded support that the Bernoulli distribution with a point mass of $1-\theta$ at 1 is the limiting distribution of X for $0 < \theta < 1$ as $n \downarrow 0$ (see, e.g., Harris 1966, p. 103). These limiting distributions coincide with the related limiting distributions of the beta distribution (see, e.g., van Dorp and Mazzuchi 2000). In other words, the flexibility of the $\mathrm{STSP}(\theta, n)$ class is comparable to that of the beta family.

Another interesting limiting distribution (relevant to the example in Section 1) can be derived using the linear transformation

$$Y = (n-1)A\left(\frac{X-\theta}{\theta}\right), \tag{17}$$

where $A > 0$ is an abritrary positive constant. From (3), (17), and $n > 1$ it follows that

$$f(y|\theta, n, A) = \begin{cases} \frac{n\theta}{(n-1)A}\left(1 + \frac{1}{(n-1)}\frac{Y}{A}\right)^{n-1} & -(n-1)A < Y \leq 0 \\ \frac{n\theta}{(n-1)A}\left(1 - \frac{1}{(n-1)}\frac{\theta Y}{(1-\theta)A}\right)^{n-1} & 0 \leq Y < \frac{(1-\theta)(n-1)A}{\theta}. \end{cases} \tag{18}$$

Letting $n \to \infty$, we have

$$f(y|\theta, n, A) \to f(y|\theta, A) = \begin{cases} \frac{\theta}{A}\exp\left(\frac{Y}{A}\right) & Y \leq 0 \\ \frac{\theta}{A}\exp(-\frac{\theta}{1-\theta}\frac{Y}{A}) & Y \geq 0. \end{cases} \tag{19}$$

Hence, $f(y|\theta, n, A)$ converges to the density $f(y|\theta, A)$ of an "asymmetric Laplace" variable (see, e.g., Johnson et al. 1994) where its probability mass is split at 0 into θ and $(1-\theta)$, regardless of the value of the parameter A. The AL distribution considered by Kozubowski and Podgórski (1999) (see (1)) simplifies to (19) by setting $\mu = 0$, $\kappa = \sqrt{1-\theta/\theta}$, $\sigma = \frac{1}{A}\sqrt{1-\theta/\theta}$.

2.5 Relative Entropy

The relative entropy (also known as *cross entropy or discrimination function*) of an absolute continuous probability density function $f(x|\Theta)$ with respect to a probability mass function $g(x)$ is defined as

$$E(f : g|\Theta) = \int \log\frac{f(x|\Theta)}{g(x)}dF(y|\Theta), \tag{20}$$

and is used as a measure for comparing information content of distributions. The term *discrimination*, reflects the fact that $E(f; g|\Theta) \geq 0$ and the equality holds if and only if $f(x|\Theta) = g(x)$ almost everywhere (see, e.g., Soofi and Retzer 2002).

We compare the information contents of STSP distributions on $[0, 1|$ to the information content of a uniform$|0, 1|$ distribution. The relative entropy of beta distributions with respect to a uniform$|0, 1|$ distribution has been studied (see, e.g., Soofi and Retzer 2002) resulting in the expression

$$\begin{aligned} E(f : g|\alpha, \beta) &= \log(\mathbb{B}(\alpha, \beta)) \\ &-(\alpha-1)(\psi(\alpha) - \psi(\alpha+\beta)) - (\beta-1)(\psi(\beta) - \psi(\alpha+\beta)) \end{aligned} \tag{21}$$

where

$$f(x|\alpha, \beta) = \frac{1}{\mathbb{B}(\alpha, \beta)} x^{\alpha-1}(1-x)^{\beta-1}, \tag{22}$$

$\alpha > 0$, $\beta > 0$, $\mathbb{B}(\alpha, \beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha+\beta)$, g is the uniform$[0, 1]$ pdf and $\psi(\cdot) = \Gamma'(\cdot)$ is the psi-function.

The relative entropy of STSP distributions with respect to a uniform$|0, 1|$ distribution, using (3) and (20), results in

$$E(f : g \,|\, \theta, n) = \log n - \frac{n-1}{n}, \tag{23}$$

where f is the STSP density given by (3) and g is as above. $E(f : g \,|\, \theta, n)$ attains its minimal value 0 when the STSP variable coincides with a uniform$|0, 1|$ variable—that is, $n = 1$. Note that the forms of (23) and (21) are similar, except that $E(f : g \,|\, \theta, n)$ is constant for fixed n regardless of the value θ. Hence, no information is added to or subtracted from the information content of a STSP distribution by varying the parameter θ, while keeping n fixed. Consequently, the relative entropy of all triangular distribution on $[0, 1]$ equals $-1/2+\log(2)$, regardless of the location of the mode θ. It is worth noting that the variance (which is intuitively related to entropy) of an $\mathrm{STSP}(\theta, n)$ variable *does* depend on θ.

3. MLE METHOD FOR TWO PARAMETER STSP DISTRIBUTION

The proposed derivation of the MLE procedure of a STSP distribution is quite instructive. Let for a random sample $\underline{X}= (X_1, \ldots, X_s)$ the order statistics be $X_{(1)} < X_{(2)} < \cdots < X_{(s)}$. By definition the likelihood for $\underline{X}$ is

$$L(\underline{X}; \theta, n) = n^s \left\{ \frac{\prod_{i=1}^{r} X_{(i)} \prod_{i=r+1}^{s}(1-X_{(i)})}{\theta^r(1-\theta)^{s-r}} \right\}^{n-1}, \tag{24}$$

where $X_{(r)} \leq \theta < X_{(r+1)}$, with $X_{(0)} \equiv 0$, $X_{(s+1)} \equiv 1$.

Theorem 2. Let $\underline{X}$= $(X_1, \ldots, X_s)$ be an iid sample from a STSP(θ, n) distribution. The MLE estimators maximizing (24) are

$$\begin{cases} \widehat{\theta} = X_{(\widehat{r})} \\ \widehat{n} = -\frac{s}{\log M(\widehat{r})}, \end{cases} \tag{25}$$

where $\widehat{r} = \arg\max_{r\in\{1,\ldots,s\}} M(r)$ and

$$M(r) = \prod_{i=1}^{r-1} \frac{X_{(i)}}{X_{(r)}} \prod_{i=r+1}^{s} \frac{1-X_{(i)}}{1-X_{(r)}}. \tag{26}$$

Proof: To maximize the likelihood (24), we set

$$\max_{n>0,\, 0\le\theta\le 1} L(\underline{X};\theta, n) = \max_{n>0}\left[n^s \cdot \widehat{M}^{n-1}\right], \tag{27}$$

where $\widehat{M}$ is given by

$$\widehat{M} = \max_{0\le\theta\le 1}\left[\frac{\prod_{i=1}^r X_{(i)} \prod_{i=r+1}^s (1-X_{(i)})}{\theta^r(1-\theta)^{s-r}}\right], \tag{28}$$

and as above $X_{(r)} \le \theta < X_{(r+1)}$, with $X_{(0)} \equiv 0$, $X_{(s+1)} \equiv 1$. Now

$$\log\left\{n^s\widehat{M}^{n-1}\right\} = (n-1)\log\widehat{M} + s\log n, \tag{29}$$

and

$$\frac{\partial}{\partial n}\log\left\{n^s\widehat{M}^{n-1}\right\} = \log\widehat{M} + \frac{s}{n}. \tag{30}$$

Equating (30) to zero yields $\widehat{n} = -(s/\log\widehat{M})$. From (30) it follows that

$$\frac{\partial}{\partial n}\log\left\{n^s\widehat{M}^{n-1}\right\} > 0 \Leftrightarrow \widehat{n} < -\frac{s}{\log\widehat{M}}. \tag{31}$$

Hence $\widehat{n}$ corresponds to a global maximum of both (29) and (27). Note that for $i < r$ it follows that $0 < X_{(i)}/\theta < 1$ and for $i > r$ it follows that $0 < (1-X_{(i)})/(1-\theta) < 1$. Hence, $0 < \widehat{M} < 1$ and thus $\widehat{n} > 0$. Using (28) we may write $\widehat{M} = \max_{r\in\{0,\ldots,s\}} H(r)$ where

$$H(r) = \max_{X_{(r)}\le\theta\le X_{(r+1)}}\left[\frac{\prod_{i=1}^r X_{(i)} \prod_{i=r+1}^s (1-X_{(i)})}{\theta^r\cdot(1-\theta)^{s-r}}\right]. \tag{32}$$

We shall discuss separately the three cases: $r \in \{1,\ldots,s-1\}$, $r = 0$ and $r = s$.

Case $r \in \{1,\ldots,s-1\}$: Here, $X_{(r)} \le \theta \le X_{(r+1)}$. The function $g(\theta) = \theta^r(1-\theta)^{s-r}$ is proportional to a *unimodal* beta density. Thus,

$$\min_{X_{(r)}\le\theta\le X_{(r+1)}} g(\theta) = \min_{\theta\in\{X_{(r)},X_{(r+1)}\}} g(\theta) \tag{33}$$

and, from (32),

$$H(r) = \max_{r'\in\{r,r+1\}} \prod_{i=1}^{r'-1} \frac{X_{(i)}}{X_{(r')}} \prod_{i=r'+1}^{s} \frac{1-X_{(i)}}{1-X_{(r')}}. \tag{34}$$

Case $r = 0$: Here $0 \le \theta \le X_{(1)}$. From (32) it follows that in this case

$$H(0) = \max_{0\le\theta\le X_{(1)}}\left[\prod_{i=1}^s \frac{1-X_{(i)}}{1-\theta}\right]. \tag{35}$$

Hence,

$$H(0) = \prod_{i=1}^s \frac{1-X_{(i)}}{1-X_{(1)}} = \prod_{i=2}^s \frac{1-X_{(i)}}{1-X_{(1)}}. \tag{36}$$

Case $r = s$: Here $X_{(s)} \le \theta \le 1$. From (32) it follows that in this case

$$H(s) = \max_{X_{(s)}\le\theta\le 1}\left[\prod_{i=1}^s \frac{X_{(i)}}{\theta}\right]. \tag{37}$$

Hence

$$H(s) = \prod_{i=1}^s \frac{X_{(i)}}{X_{(s)}} = \prod_{i=1}^{s-1} \frac{X_{(i)}}{X_{(s)}}. \tag{38}$$

From (34), (36), and (38) we obtain that $\widehat{M} = \max_{r\in\{1,\ldots,s\}} M(r)$ where

$$M(r) = \prod_{i=1}^{r-1} \frac{X_{(i)}}{X_{(r)}} \prod_{i=r+1}^{s} \frac{1-X_{(i)}}{1-X_{(r)}}. \tag{39}$$

Note that, $\widehat{M}$ is attained at $\widehat{\theta} = X_{(\widehat{r})}$ where $\widehat{r} = \arg\max_{r\in\{1,\ldots,s\}} M(r)$.

The estimates given in (25) are quite intuitive. In particular the estimator of the parameter θ (governing location and skewness of the distribution) is a specific order statistic. We note, in passing, that the approach for determining the MLE estimate $\widehat{\theta}$ for a STSP(θ,n) distribution is similar (though simplified) to the approach for determining the MLE estimate $\widehat{\theta}$ for a $\mathrm{triang}(0,\theta,1)$ distribution (see Johnson and Kotz 1999).

We shall illustrate the MLE procedure for an $\mathrm{STSP}(\theta, n)$ distribution using the following fictitious order statistics

$$\begin{aligned}(X_{(1)},\ldots,X_{(10)}) = {} & (0.340, 0.395, 0.413, 0.420, 0.423, \\ & 0.429, 0.465, 0.513, 0.564, 0.588).\end{aligned} \tag{40}$$

Consider the matrix $A = [a_{i,r}]$ where

$$a_{i,r} = \begin{cases} \frac{X_{(i)}}{X_{(r)}} & i < r \\ \frac{1-X_{(i)}}{1-X_{(r)}} & i \ge r. \end{cases} \tag{41}$$

Table 1 summarizes the calculation of the matrix A for order statistics given by (40). The last row in the table contains the products of the matrix elements in the rth column which are equal to the values of $M(r)$ given by (39). Numerical calculations yield

$$\begin{aligned} \widehat{M} &= 0.304; \quad \widehat{r} = 5; \\ \widehat{\theta} &= X_{(\widehat{r})} = 0.423; \\ \widehat{n} &= \frac{-10}{\log(\widehat{M})} = 8.399. \end{aligned} \tag{42}$$

Table 1. Example of MLE Estimation for STSP(θ, n)

i	$X_{(i)}$ \ r: $X_{(r)}$	1: 0.340	2: 0.395	3: 0.413	4: 0.420	5: **0.423**	6: 0.429	7: 0.465	8: 0.513	9: 0.564	10: 0.588
1	0.340	1.000	0.861	0.823	0.810	0.804	0.793	0.731	0.663	0.603	0.578
2	0.395	0.917	1.000	0.956	0.940	0.934	0.921	0.849	0.770	0.700	0.672
3	0.413	0.889	0.970	1.000	0.983	0.976	0.963	0.888	0.805	0.732	0.702
4	0.420	0.879	0.959	0.988	1.000	0.993	0.979	0.903	0.819	0.745	0.714
5	0.423	0.874	0.954	0.983	0.995	1.000	0.986	0.910	0.825	0.750	0.719
6	0.429	0.865	0.944	0.973	0.984	0.990	1.000	0.923	0.836	0.761	0.730
7	0.465	0.811	0.884	0.911	0.922	0.927	0.937	1.000	0.906	0.824	0.791
8	0.513	0.738	0.805	0.830	0.840	0.844	0.853	0.910	1.000	0.910	0.872
9	0.564	0.661	0.721	0.743	0.752	0.756	0.764	0.815	0.895	1.000	0.959
10	0.588	0.624	0.681	0.702	0.710	0.714	0.722	0.770	0.846	0.945	1.000
	M(*r*)	0.134	0.252	0.293	0.303	**0.304**	0.299	0.239	0.159	0.093	0.068

Maximum likelihood estimation for $X \sim \mathrm{STSP}(\theta,\, n)$ distribution can be modified to two parameter maximum likelihood estimation for $Z \sim \mathrm{TSP}(a, m, b, n)$, where $Z = (b-a)X + a$, the parameters a and b are fixed and the parameter $m = (b-a)\theta + a$. We have for the pdf of Z

$$f(z|a,m,b,n) = \begin{cases} \frac{n}{(b-a)}\left(\frac{z-a}{m-a}\right)^{n-1} & a < z \leq m \\ \frac{n}{(b-a)}\left(\frac{b-z}{b-m}\right)^{n-1} & m \leq z \leq b. \end{cases} \quad (43)$$

The maximum likelihood estimates for the parameters m and n in (43) using order statistics $(Z_{(1)}, \ldots, Z_{(s)})$ are

$$\begin{cases} \widehat{m}(a,b) = Z_{(\widehat{r}(a,b))} \\ \widehat{n}(a,b) = -\frac{s}{\log M(a,b,\widehat{r}(a,b))}, \end{cases} \quad (44)$$

where, as above,

$$\widehat{r}(a,b) = \arg \max_{r \in \{1,\ldots,s\}} M(a,b,r), \quad (45)$$

and

$$M(a,b,r) = \prod_{i=1}^{r-1} \frac{Z_{(i)} - a}{Z_{(r)} - a} \prod_{i=r+1}^{s} \frac{b - Z_{(i)}}{b - Z_{(r)}}. \quad (46)$$

The MLE's in (44) were used in the example of Section 1, with fixed $a = -0.02$ and $b = 0.02$, yielding

$$\widehat{m}(-0.02, 0.02) = 0.0,\ \widehat{n}(-0.02, 0.02) = 7.6365 \quad (47)$$

The authors have also studied four parameter MLE estimation for the $\mathrm{TSP}(a, m, b, n)$ distribution (see (43)) which will hopefully be presented in a follow-up article.

4. MOMENT RATIO DIAGRAMS

Moment ratio plots popularized for Pearson-type distributions by Elderton and Johnson (1969) seem to provide a useful visual (graphical) assessment of the skewness (asymmetry) and elusive kurtosis (peakedness) inherent in any particular family of asymmetric distributions. The classical form of the diagram shows the values of the ratios

$$\beta_1 = \frac{E^2[(X-E[X])^3]}{E^3[(X-E[X])^2]} = \frac{\mu_3^2}{\mu_2^3},$$

$$\beta_2 = \frac{E[(X-E[X])^4]}{E^2[(X-E[X])^2]} = \frac{\mu_4}{\mu_2^2} \quad (48)$$

with β_1 as abscissa and β_2 as ordinate. This diagram suffers from the defect that the sign of μ_3 (indicating left skewness or right skewness) is lost. A moment ratio diagram that retains this information is a plot with $\sqrt{\beta_1}$ as abscissa and β_2 as ordinate, with the convention that $\sqrt{\beta_1}$ retains the sign of μ_3 (see, e.g., Kotz and Johnson 1985).

Values for $\sqrt{\beta_1}$ and β_2 for STSP distributions can be calculated using (5) and the relationship between central moments μ_k, $k = 2, \ldots, 4$. and the moments around the origin $\mu'_k = E|X^k|$, $k = 1, \ldots, 4$,

$$\begin{cases} \mu_2 = \mu'_2 - {\mu'_1}^2 \\ \mu_3 = \mu'_3 - 3\mu'_2\mu'_1 + 2{\mu'_1}^3 \\ \mu_2 = \mu'_4 - 4\mu'_3\mu'_1 + 6\mu'_2{\mu'_1}^2 - 3{\mu'_1}^4 \end{cases} \quad (49)$$

(see, e.g., Stuart and Ord 1994). An explicit form of $\sqrt{\beta_1}$ and β_2 for STSP distributions results in cumbersome and not very informative expressions and have therefore been omitted. Figure 5 displays the moment ratio diagram coverage for the STSP family restricted to a parameter range of

$$0.1 \leq n \leq 25, 0 \leq \theta \leq 1. \quad (50)$$

The range indicated by (50) is a plausible range for practical purposes and includes unimodal forms $(0 \leq \theta \leq 1,\, n > 1)$, U-shaped forms $(0 \leq \theta \leq 1,\, n < 1)$ as well as the uniform distribution $(n = 1)$, triangular distributions $(n = 2)$, J-shaped power function distributions $(\theta = 1,\, n > 0)$ and their J-shaped reflection $(\theta = 0,\, n > 0)$. Figure 5 also shows the effect of the parameters $(\theta,\, n)$ on $(\sqrt{\beta_1}, \beta_2)$ for specific examples of these cases indicated by solid lines in the moment ratio diagram. The shaded region in the moment ratio diagram is called the *infeasible region* since for all distributions

$$\beta_2 \geq (\sqrt{\beta_1})^2 + 1, \quad (51)$$

(see, e.g., Kotz and Johnson 1985). The horizontally (vertically) hatched area indicates the coverage of $(\sqrt{\beta_1}, \beta_2)$ for unimodal (U-shaped) STSP distributions. The only J-shaped members in the STSP family are the power function distribution ($\theta = 1$) and its reflection ($\theta = 0$) indicated by solid lines in Figure 5.

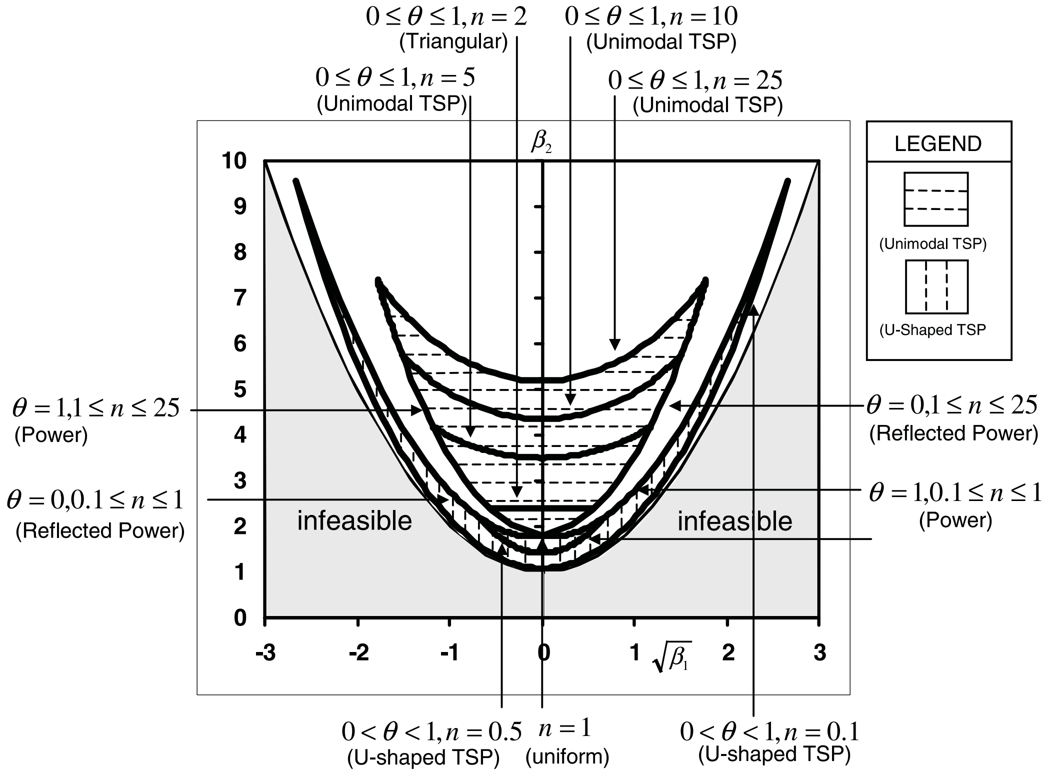

Figure 5. ($\sqrt{\beta_1}$, β_2) moment ratio diagram for STSP distributions with parameter range $0 \le \theta \le 1$,$0.1 \le n \le 25$.

For comparison purposes, we have generated a moment ratio diagram for beta densities (see (22)) using expressions for the moments of the beta distribution (see, e.g., Johnson et al. 1994). From (3) and (22) it follows that for $\theta = 1$ $(\theta = 0)$ and $\beta = 1$ $(\alpha = 1)$ the STSP density and beta density both coincide with the density of a (reflected) power function distribution. Hence, a comparable parameter range for the parameters α and β in (22) follows from (50) to be

$$0.1 \le \alpha \le 25,\ 0.1 \le \beta \le 25. \tag{52}$$

Figure 6 displays the moment ratio diagram for beta densities restricted to (52). The range indicated by (52) includes all forms of the beta density; that is, unimodal, J-shaped and U-shaped. The coverage area for $(\sqrt{\beta_1}, \beta_2)$ for unimodal, U-shaped, and J-shaped beta densities are indicated in Figure 6 by horizontal, vertical, and cross hatched areas, respectively. The effect of the parameters α and β in (22) on $(\sqrt{\beta_1}, \beta_2)$ are indicated by solid lines for those cases of beta densities that identify the boundaries of the hatched areas. The top boundary (not solid) was generated by interpolation using moment ratio curves for the cases of the form $\alpha = c,\ 1 \le \beta \le 25$ and $\beta = c, 1 \le \alpha \le 25$ with $c \in \{0.5, 2, 5, 10\}$. These curves are not included in Figure 6 to amplify identification of the hatched areas.

The comparison of Figures 5 and 6 is illuminating. First, Figure 6 shows that, in terms of moment ratio coverage, the beta family is richer than the STSP family when restricted to J-shaped forms. The only J-shaped STSP distributions—that is, the power function distribution and its reflection—are represented within the beta family as indicated by the only common solid lines in Figures 5 and 6. The intersection of these lines identifies the only other common member of the STSP family and beta family—that is, the uniform$|0, 1|$ distribution. Second, the coverage areas associated with U-shaped forms in Figure 5 and Figure 6 is comparable in size indicating similar flexibility between the STSP family and beta family when restricted to these forms. Finally, possibly most importantly, the coverage area of the beta family restricted to unimodal forms in Figure 6 is completely contained *within* the coverage area of the STSP family restricted to unimodal forms. The latter observation indicates greater flexibility by the STSP family than the beta family when modeling unimodal phenomena where the mode is not at a support boundary and smooth behavior of the density function at its mode is not a requirement.

The STSP distribution may be considered as an alternative to the beta distribution when sample estimates for skewness $\sqrt{\beta_1}$ and kurtosis β_2 fall outside the coverage area indicated in Figure 6. For example, from the data in Figure 1(b) we estimate $\sqrt{\widehat{\beta}_1} = -0.05451$ and $\widehat{\beta}_2 = 6.07494$ outside the coverage area in Figure 6. For the beta distribution with parameters $\widehat{\alpha} = 18.227$, $\widehat{\beta} = 18.366$ associated with Figure 3 we get skewness $\sqrt{\beta_1} = 0.00242$ and kurtosis $\beta_2 = 2.84847$. For the STSP distribution with parameters $\widehat{\theta} = 0.5$, $\widehat{n} = 7.6365$ associated with Figure 3 we get skewness $\sqrt{\beta_1} = 0.0$ and kurtosis $\beta_2 = 4.03448$. Although the estimated STSP distribution does not totally capture kurtosis $\widehat{\beta}_2$ observed in the data, the STSP family provides a better fit in terms of $\widehat{\beta}_2$ (and $\sqrt{\widehat{\beta}_1}$) than the estimated beta distribution. It is worth noting that values for kurtosis

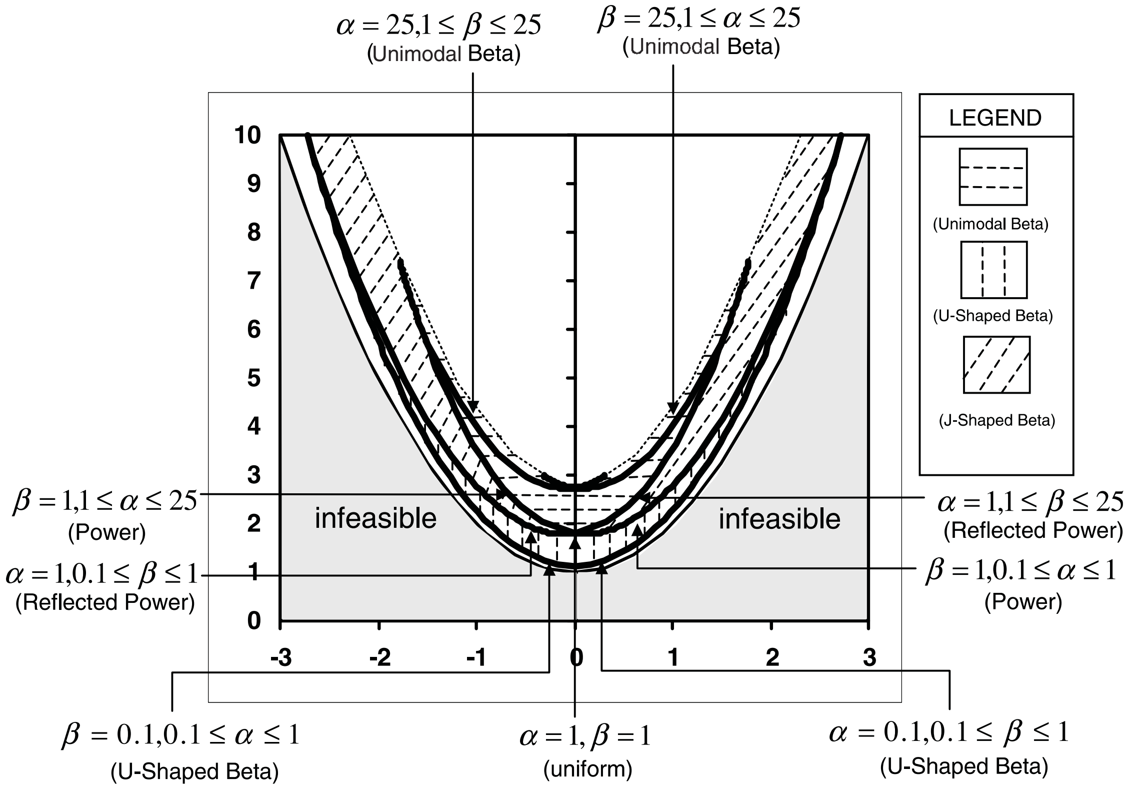

Figure 6. ($\sqrt{\beta_1}$, β_2) moment ratio diagram for Beta distributions with parameter range $0.1 \le \alpha \le 25$, $0.1 \le \beta \le 25$.

β_2 for symmetric unimodal beta distributions with parameters restricted to $1 < \alpha < 25, \beta = \alpha$ are strictly less than 2.88679 and values for kurtosis β_2 for symmetric unimodal STSP distributions with parameters restricted to $\theta = 0.5$, $3.0745 < n < 25$ are strictly larger than 2.88679.

Finally, from Figure 5 it may be observed that the parameter θ of an STSP density primarily affects skewness $\sqrt{\beta_1}$ which reemphasizes the role of θ as a location parameter. The parameters α and β of the beta distribution affect both skewness $\sqrt{\beta_1}$ and kurtosis β_2 in a similar manner and henceforth do not allow for such an interpretation.

5. CONCLUDING REMARKS

A new family of distributions is proposed which possesses attractive properties, especially those related to the meaning of the parameters, the structure of its expected value as a function of parameters, a closed form expression for its cdf, a MLE procedure involving only elementary functions and a transparent form of its entropy function. These properties are not shared by the beta family. Similar to the beta family, the new family allows for U-shaped, J-shaped and unimodal forms. The density of the new family is non-smooth (non-differentiable) at θ (cf. (3)). This "drawback" seems to be of a lesser importance in modern engineering, economic, and financial data. It is worth noting that for some 65 years after Karl Pearson's death no serious attempts have been made to offer an alternative to the beta density of comparable flexibility, which shows the prominence of Karl Pearson's discovery of the beta distribution.

For parameter values in the range $1 \le n \le 3$ (see (3)) the new family displays unimodal forms with a modest peak at θ and complements the beta family when smooth behavior at the mode is not a crucial requirement while any of the above properties are desirable. For parameter values of $n > 3$ the family adds to existing modeling capabilities of unimodal phenomena on a bounded domain in particular when a peak in data is observed (see, e.g., Figure 3). For parameters values in the range of $0 \le n < 1$, the family primarily exhibits U-shaped forms similarly to the beta distribution. The only J-shaped distribution within the new family is the power function distribution ($\theta = 1$) (cf. (3)) and its reflection ($\theta = 0$) which are also shared by the beta family. Hence, the beta distribution enjoys greater flexibility among the J-shaped distributions. Summarizing, the differences between the new family and the beta family are quite similar to those between the Laplace family (which is becoming more popular) and the normal family (see, e.g., Kotz, Kozubowski, and Podgórski (2001), but restricted to a bounded support.

The analysis of the two parameter two-sided power distribution can be extended to the four parameter case involving the boundary parameters. Although the MLE procedure is more delicate in this case, it is algorithmically straightforward using modern computational facilities. It is our hope that the introduction of the proposed distribution into statistical theory and practice will contribute to the basic goals of applied statistical work: reaching the point when accumulation of data on a specific issue is directly followed by an understanding of the meaning of its parameters.

[Received October 2000. Revised December 2001.]

REFERENCES

Elderton, W. P., and Johnson, N. L. (1969), *Systems of Frequency Curves*, London: Cambridge University Press.

Harris, B. (1966), *Theory of Probability*, Reading, MA: Addison Wesley.

Johnson, D. (1997), "The Triangular Distribution as a Proxy for the Beta Distribution in Risk Analysis," *The Statistician*, 46, 387–398.

Johnson, N. L, and Kotz, S. (1999), "Non-Smooth Sailing or Triangular Distributions Revisited after Some 50 Years," *The Statistician*, 48, 179–187.

Johnson, N. L., Kotz, S., and Balakrishnan, N. (1994), *Continuous Univariate Distributions–1* (2nd ed.), New York: Wiley.

——— (1995), *Continuous Univariate Distributions - 2*, (2nd ed.), Wiley, New York.

Klein, G. E. (1993), "The Sensitivity of Cash-Flow Analysis to the Choice of Statistical Model for Interest Rate Changes" (with discussions), TSA XLV, 79–186.

Kotz, S., and Johnson N. L., (eds.) (1985), "Moment Ratio Diagrams," in *Encyclopedia of Statistical Sciences* (vol. 5), New York: Wiley, pp. 602–604.

Kotz, S., Kozubowski, T. J., and Podgórski, K. (2001), *The Laplace Distribution and Generalizations, A Revisit with Applications to Communications, Economics, Engineering and Finance*, Boston: Birkhäuser.

Kozubowski, T. J., and Podgórski, K. (1999), "A Class of Assymetric Distributions," in *Actuarial Research Clearing House* (vol. 1), pp. 113–134.

Leunberger, D. G. (1998), *Investment Science*, New York: Oxford University Press.

Mielke, P. W., Jr. (1975), "Convenient Beta Distribution Likelihood Techniques for Describing and Comparing Meteorological Data," *Journal of Applied Meteorology*, 14, 985–990.

Nadarajah, S. (1999), "A Polynomial Model for Bivariate Extreme Value Distributions," *Statistics & Probability Letters*, 42, 15–25.

Soofi, E. S., and Retzer, J. J. (2002), "Information Indices: Unification and Applications," *Journal of Econometrics*, 107, 17–40.

Stuart, A., and Ord, J. K. (1994), "Kendall's Advanced Theory of Statistics," in *Distribution Theory* (vol. 1), New York: Wiley, 74–75.

Van Dorp, J. R., and Mazzuchi, T. A. (2000), "Solving for the Parameters of a Beta distribution under Two Quantile Constraints," *Journal of Statistical Computation and Simulation*, 67, 189–201.

Williams, T. M. (1992), "Practical Use of Distributions in Network Analysis," *Journal of Operations Research Society*, 43, 265–270.